

#5 attachment
09/670568

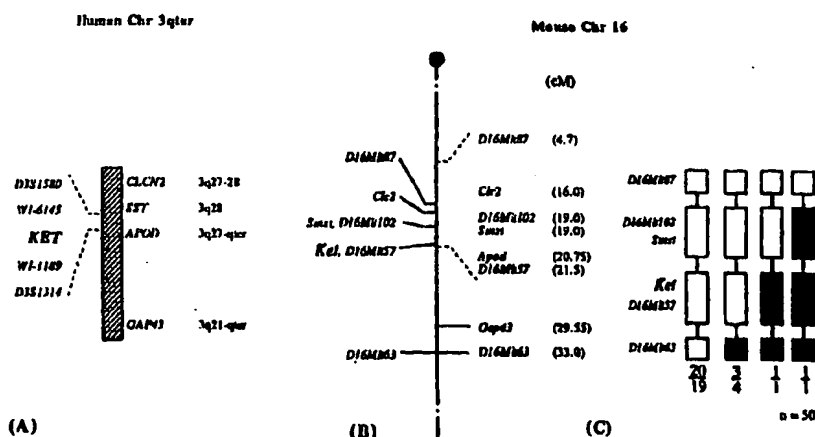


PCT WELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales Büro
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 6 : C12N 15/12, 5/10, C07K 14/47, C12Q 1/68, G01N 33/574		A2	(11) Internationale Veröffentlichungsnummer: WO 99/61610
		(43) Internationales Veröffentlichungsdatum:	2. Dezember 1999 (02.12.99)
(21) Internationales Aktenzeichen: PCT/DE99/01557		(81) Bestimmungsstaaten: CA, JP, US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) Internationales Anmeldedatum: 25. Mai 1999 (25.05.99)		Veröffentlicht <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i>	
(30) Prioritätsdaten: 198 22 985.2 25. Mai 1998 (25.05.98) DE			
(71) Anmelder (für alle Bestimmungsstaaten ausser US): FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V. [DE/DE]; Leonrodstrasse 54, D-80636 München (DE).			
(72) Erfinder; und (75) Erfinder/Anmelder (nur für US): PAUL, Dieter [DE/DE]; Alsterblick 24, D-22397 Hamburg (DE). AUGUSTIN, Martin [DE/DE]; Stockenstrasse 15, D-53113 Bonn (DE). SCHMALE, Hartwig [DE/DE]; Goldkäferweg 62, D-22523 Hamburg (DE). BAMBERGER, Casimir [DE/DE]; Ernst-Thälmann-Platz 3, D-20251 Hamburg (DE).			
(74) Anwalt: BAUMBACH, Fritz; Robert-Rössle-Strasse 10, D-13125 Berlin (DE).			

(54) Title: **TUMOUR SUPPRESSOR GENES OF THE p53 FAMILY**

(54) Bezeichnung: **TUMORSUPPRESSORGENE DER p53-FAMILIE**



(57) Abstract

The invention relates to novel tumour suppressor genes of the p53 family, to polypeptides which code them and to their use. It preferably relates to nucleic acids which code KET, especially those of rats, humans and mice.

(57) Zusammenfassung

Die Erfindung betrifft neue Tumorsuppressorgene der p53-Familie, Polypeptide die sie kodieren, sowie ihre Verwendung. Bevorzugt betrifft sie KET-kodierende Nukleinsäuren, insbesondere der Ratte, des Menschen und der Maus.

LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidshan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauretanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LI	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		
EE	Estland						

Tumorsuppressorgene der p53-Familie

Die Erfindung betrifft neue Tumorsuppressorgene der p53-Familie, Polypeptide, die sie kodieren sowie ihre Verwendung. Bevorzugt betrifft sie KET-kodierende Nukleinsäuren, insbesondere der Ratte, des Menschen und der Maus.

Gene, die in der Tumorigenese eine maßgebliche Rolle spielen, können aufgrund ihrer funktionellen Wirkweise grob klassifiziert werden. Führt eine Gain-of-function-Mutation zu einem Allel, das auf die Tumorigenese aktivierend wirkt, so wird das betroffene Gen als Oncogen bezeichnet. Ist eine Loss-of-function-Mutation auf beiden Allele nötig (inaktivierend), um tumorigene Veränderungen möglich werden zu lassen, spricht man von einem Tumorsuppressor-Gen. Das prominenteste und meist untersuchte Tumorsuppressor-Gen kodiert für den nukleären Transkriptionsfaktor p53 (über 2000 Medline Einträge im vergangenen Jahr; NCBI-Datenbank) mit der Hauptfunktion in der Kontrolle des Zell-Zyklus und der Apoptose (Levine 1997). p53 liegt in humanen Tumoren zu über 50 % mutiert vor (Hollstein et al. 1991), vererbte p53 Mutationen führen zu einer erhöhten Tumorfrequenz bei den Trägern des defekten Allels (Evans und Lozano 1997). Derzeit existieren vier p53-knock-out Mauslinien, die von mehreren Arbeitsgruppen unabhängig voneinander hergestellt wurden und als Tiermodelle für den Menschen dienen. p53-defiziente Tiere (Genotyp: +/- und -/-) zeigen schon im frühen Alter eine vermehrte Tumorraten (Donehower et al. 1992, Harvey et al. 1993b). Die Embryo- und Organogenese verläuft jedoch im allgemeinen unauffällig, so daß p53 defiziente Tiere nicht von ihren +/- Wildtypgeschwistern zu unterscheiden sind (Donehower et al. 1992). Allerdings zeigen einige p53 -/- Embryonen am Tag 13.5 der Embryogenese einen charakteristischen Defekt in der Morphologie des Kopfbereiches, eine Exencephalie, bei der der Verschluss des Neuralrohres in Vorder und Mittelhirn nicht erfolgt. Diese Mißbildung tritt bei 16 % der homozygot-defizienten CV 129 Embryonen auf, während nur 8 % von CV129 X C57BL/6- p53 -/- Hybridembryonen betroffen sind. Interessanterweise variiert auch die Tumorzinzidenz im CV129 Hintergrund und CV129 X C57BL/6 Hybridhintergrund. Tumoren entwickeln sich generell schneller in CV 129 Mäusen als in CV 129 X C57BL/6-Hybriden; zusätzlich kommt es in CV 129 Tieren vermehrt zu Teratomen (Donehower et al. 1995, Harvey et al. 1993a). Solche Unterschiede

können nur durch die jeweils unterschiedliche Konstitution des genetischen Hintergrundes erklärt werden. Das zeigt im unterschiedlichen genetischen Hintergrund das Vorliegen differenzierter Kompensationseffizienz gegenüber dem p53-Verlust und die Existenz verwandter Gen-Produkte, welche die Funktion von p53 während der Embryogenese und Cancerogenese verrichten.

Der Erfindung lag deshalb die Aufgabe zugrunde, entsprechende Gene bereitzustellen, die für Proteine kodieren, welche bei der Kontrolle des Zell-Zyklus und der Apoptose eine Rolle spielen.

Die Erfindung basiert auf der Erkenntnis, daß das Protein KET mit so bemerkenswerter Homologie in seiner Aminosäuresequenz zu p53 gefunden wurde, daß es mit p53 in einer p53-Familie zusammengefaßt werden kann. Wie p53 besitzt KET eine Transaktivierungs-, eine DNA-Bindungs- und eine Oligomerisierungsdomäne. Der höchste Grad an Homologie ist in der DNA-Bindungs-Domäne zu finden. Er beträgt zwischen KET und p53 75%. Die Isolierung der kodierenden cDNAs erfolgte aus der Ratte (SEQ ID No. 1).

Erfindungsgemäß wurde die menschliche KET cDNA (SEQ ID No. 2) kloniert; ein Alignment der abgeleiteten Aminosäuresequenz (SEQ ID No 3) mit der aus Ratte und den Sequenzen von humanem p53 und p73 ist in Abb. 1 gezeigt. Die KET-Aminosäure-Sequenz aus der Ratte zeigt eine Homologie von 98 % zu der des Menschen.

Es wurde eine chromosomale Lokalisation des Gens auf dem Chromosom 3q des Menschen und 16 der Maus (vgl. Abb. 2 Genetische Kartierung) gefunden, woraus sich auf eine Funktion von KET als Tumorsuppressor rückschließen läßt. Interessanterweise kartiert das *Ket*-Gen der Maus in einen Bereich, der in frühen Stadien der Pancreaskanzerogenese deletiert ist und vermutlich einen Suppressor der Angiogenese, Loh2 (Gensymbol: Loh2), beinhaltet, für den Ket somit einen Kandidaten darstellt.

Gemäß der Erfindung wurde festgestellt, daß das KET-Protein bei der Tumorsuppression beteiligt ist. Von speziellem Interesse waren vor allem solche Tumoren, in denen bisher keine Veränderungen des p53 Wildtypallels beschrieben wurden. Die chromosomale Lokalisation der verantwortlichen

Tumorsuppressorgene kann durch cytogenetische Analysen, die Loss of heterozygosity (LOH)-Bereiche identifizieren, vorausgesagt werden. Es wurde nachgewiesen, daß das *KET/Ket*-Gen bei Mensch oder Maus in solche LOH-Regionen kartiert.

Erfindungsgemäß erfolgte eine Kartierung des *KET/Ket*-Gens bei Mensch und Maus mit flankierenden Markern (Abb.2). Zur präzisen chromosomalen Lokalisation beim Menschen wurden Bestrahlungshybride (Radiation Hybrids; GeneBridge 4 Panel, Research Genetics) eingesetzt, die Kartierung bei der Maus erfolgte in einer *M. musculus* X *M. spretus* Rückkreuzungsgeneration. Das *Ket*-Gen kartiert zwischen das Somatostatin-Gen und das Apolipoprotein D Gen auf Chr. 3q des Menschen. Dieselbe Genreihenfolge wurde auf Chr. 16 der Maus bestätigt. A (links) chromosomaler Abschnitt des humanen Chromosoms 3q mit der Position des *KET*-Locus; B (mitte) Position des *Ket*-Locus auf Chr. 16 der Maus. C (rechts) *Ket*-Haplotypen und Markergene auf Chr. 16. Jede Säule repräsentiert zwei Haplotypen von Chr. 16, die Anzahl der Backcross-Individuen ist unten angegeben (obere Zahl: gefüllte Quadrate/Rechtecke SEG/1 Allel; leere Quadrate/Rechtecke C57BL6J Allel; untere Zahl: das Gegenteil)
Gensymbole: *CKCN2/Clc2* – Chlorid-Channel 2; *SST/Smst* – Somatostatin; *KET/Ket* – p53 verwandtes Protein *KET*; *APOD/Apod* – Apolipoprotein D; *GAP43/Gap43* – Wachstum beeinflussendes Protein 43.

Gegenstand der Erfindung sind deshalb die *KET*-Nukleinsäuren, vorzugsweise *KET*-cDNA der Ratte, des Menschen und der Maus sowie deren Fragmente, Varianten und Mutationen, vorzugsweise die SEQ ID No.1 (*KET*-cDNA der Ratte) und die SEQ ID No. 2 (humane *KET*-cDNA).

Fragmente, Varianten und Mutationen sind durch Basenaustausche gekennzeichnet. Weiterhin können auch alle T durch U ersetzt sein (Ribonukleinsäure).

Ferner sind Gegenstand der Erfindung die Polypeptide für die die cDNAs kodieren, vorzugsweise SEQ ID No. 3, und deren an einer oder mehreren Stellen durch Austausch von Aminosäuren geänderte Strukturen.

Die Herstellung erfolgt nach an sich bekannten Verfahren, wie z. B. durch Isolierung und Sequenzierung aus cDNA-Bibliotheken.

Desweiteren betrifft die Erfindung die Verwendung der KET-Nukleinsäuren und Polypeptide als Ausgangsbasis zur Entwicklung spezifischer und wirkungsvoller Cancerostatika. Sie werden zum Aufbau von Genen und Vektoren eingesetzt, die die Basis für die Entwicklung dieser pharmazeutisch relevanten Substanzen darstellen.

Außerdem werden sie zur Entwicklung diagnostischer Kits eingesetzt, so z.B. zur Vorhersage eines Krebsrisikos. Gegenstand sind demzufolge auch diagnostische Testkits.

Weiterhin erfolgte die Charakterisierung von genomischen Ket-Klonen des Menschen und der Maus durch die differentielle Darstellung von intron-umspannenden PCR-Fragmenten. Diese PCR-Tests wurden auch für die Identifizierung von Ket-positiven BACs (Bacterial artificial chromosome) in einer genomischen BAC-Bibliothek (Genome Systems Inc) verwendet. Diese BACs enthalten DNA aus dem CV 129/J Mausstamm, so daß Subfragmente des BAC-Klones direkt zur Konstruktion des Ket-Targeting-Vektors eingesetzt werden. Auf den bisher identifizierten BACs befanden sich auch die ersten 5'-Exons. Aus der full-length KET-cDNA Sequenz des Menschen und der Ratte wurden Primersequenzen abgeleitet, die zu Maus Exon 1 spezifischen PCR-Tests führen.

Darüber hinaus wurden Ket-defiziente Mäuse hergestellt.

Die Herstellung von Mäusen, die Nullallele des Ket-Gens tragen, erforderte vier aufeinanderfolgende Prozesse:

- Isolierung und Charakterisierung eines geeigneten Abschnittes des Zielgenes.
- Klonierung eines Targeting-Vektors, in dem der offene Leserahmen des Zielgenes durch die Integration eines Selektionsmarkers (vollständige Transkriptionseinheit für das Neomycin-Resistenz-Gen) zerstört ist.
- Homologe Integration des Rekombinationskonstruktes in das Genom von embryonalen Stammzellen und anschließende Selektion auf die Antibiotika-Resistenz.

- Injektion von ESCs in Blastocysten (oder Kokultur mit Morulae) und anschließender Uterustransfer.

Gegenstand sind auch Targeting-Vektoren. In einer Ausführungsvariante wurde der bereits vorliegende BAC nach Verdau mit geeigneten Restriktionsendonukleasen in pBluescript oder pUC subkloniert (optimale Größe der Subklone 5-10 kb). Über weitere Restriktionskartierung und STS-Mapping (Sequence Tagged Sites) mittels PCR und/oder Southern blotting wurde ein Subklon-Kontig erstellt, das als eine Feinkartierung des 5'-Genbereichs angesehen werden kann. Hierzu wurde die cDNA-Information aus Mensch und Ratte verwendet. Generell kann die Unterbrechung des offenen Leserahmens an zwei Stellen erfolgen: Es wurde in einem Fall die ersten translatierten Exons und im anderen die gesamte putative DNA-Bindungs-Domäne ausgeschaltet. Um die gewebsspezifische Expression von *Ket* während der Embryo- und Organogenese in Chimären und *Ket*-defizienten Mäusen zu verfolgen, wurde zusätzlich eine β -Galactosidase-Kassette so einkloniert, daß sie der Kontrolle des *Ket*-Promoters unterliegt. Generell waren zwei Formen des gezielten Gentargeting möglich: Bei Verwendung eines Insertionsvektors integrierte der komplette Vektor (ein Crossover-Ereignis war erforderlich), bei dem gebräuchlicheren Replacementvektor integrierte lediglich ein Teil, der von der Wahl intragener Restriktionsschnittstellen abhängig war (zwei Crossover-Ereignisse erforderlich).

Genausschaltung in embryonalen Stammzellen (ESCs)

Wie schon oben aufgeführt, ist die verwendete genomische BAC-Bibliothek CV 129-Ursprungs. Aus diesem Mausstamm wurden auch die meisten gebräuchlichen ESCs isoliert. Der Vorteil der Verwendung von isogenem Material liegt in der höheren Wahrscheinlichkeit zur homologen Rekombination. ESCs wurden nach Transfektion (Elektroporation) durch G418-Selektion überprüft. Im Falle des Replacement-Vektors wurde zusätzlich eine Vektor-interne Thymidinkinase-Kassette zur negativ-Selektion (Gancyclovir) bei nicht homologer Integration genutzt.

Eine erfolgreiche homologe Rekombination wurde über DNA-Analysen (Southern-Blot) geprüft.

Durch die Injektion von so geänderten ESCs in Blastocysten wurden auch Embryonalchimären hergestellt und diagnostiziert.

Für die Blastocysteninjektion oder Morula-Aggregation wurden nur genotypisierte ESCs verwendet. Chimäre Präimplantations-Embryonen wurden in die Uterushörner von scheinschwangeren Rezipientenmäusen übertragen. Chimären wurden in Testverpaarungen auf eine erfolgte Keimbahntransmission von ESC-Abkömmlingen überprüft. F1/F2 Nachkommen von Chimären, die das *Ket* Null-Allel entweder hetero- oder homozygot tragen, wurden mittels Southern-Blot- oder PCR-Analysen genotypisiert.

Weiterhin ist die Erfindung durch ein Beispiel und ein Sequenzprotokoll näher erläutert.

Beispiel

Gewinnung der humanen KET-cDNA (SEQ ID No. 2)

Für die Gewinnung von humaner KET-cDNA wurden 1×10^5 Klone einer menschlichen Skelettmuskel-cDNA-Bibliothek (Stratagene) mit Proben überprüft, die einer Ratten-KET-cDNA entstammten (Schmale und Bamberger, 1997). Ein einziger positiver Klon, hu41m, wurde gewonnen und das Insert von 3226 bp wurde unter Verwendung vektorspezifischer und interner Primer - in zwei Richtungen sequenziert. Das Insert enthielt einen offenen Leserahmen von 1360 bp, homolog zum N-Terminus der Ratten-KET-Sequenz, die Kodiersequenz war jedoch nach dem QQHQHLLQ-Motiv an Position 448 durch eine unbekannte Sequenz unterbrochen. Die Überprüfung von 6×10^5 Klonen einer menschlichen Keratinocyten-cDNA-Bibliothek (Clontech) mit einer Probe, die vom 3'-Ende von hu41m stammte, ergab zwei übereinandergreifende Klone, hu6k und hu10k, die mit einem Teil des cDNA-Klons hu41 identisch waren und die Sequenz zum 3'-Ende hin verlängerten. Um 3'-Endsequenzen zu erhalten, wurde die EST-Datenbank mit dem nichttranslatierten 3'-Bereich des Ratten-KET-Klons durchforscht. Nach Feststellung mehrerer homologer EST-Klone wurden zwei davon (I.M.A.G.E. Consortium Klon ID 149663 und 137665) vollständig sequenziert. Für die Amplifikation und direkte Sequenzierung eines 1,2 Kb-Fragmentes aus der menschlichen Haut-cDNA wurden PCR-Primer gemäß dem 3'-Ende des cDNA-Klons hu10k und dem 5'-Ende des EST-Klons 149663 verwendet. Die vollständige cDNA enthält 4846 bp, einschließlich 27 bp des höchstwahrscheinlich verkürzten nichttranslatierten 5'-Bereichs und 2776 bp des nichttranslatierten 3'-Bereichs. Um die benachbarte Anordnung der aus

verschiedenen Quellen gewonnenen Sequenzen zu demonstrieren, wurden PCR-Primer, die gemäß der Translationsstart- und -stoppkodons positioniert waren, für die Amplifikation der vollständigen Proteinkodierungssequenzen des KET von menschlicher Haut-cDNA verwendet. Die cDNA enthält einen offenen Leserahmen, der für 680 Aminosäuren kodiert (Abb. 1). Dem vermutlichen Start Methionin geht ein Translationsstoppkodon (nicht dargestellt) im Raster voran. Ein Vergleich der Aminosäuresequenzen der KET von Menschen und Ratten zeigt eine 98 %-ige Identität (Abb. 1). Diese beachtliche interspezifische Konservierung von KET-Proteinen erstreckt sich über die gesamte Moleküllänge. Sie ist im Mittelteil, der den DNA-Bindebereich enthält, sogar noch ausgeprägter; 248 Aminosäuren sind völlig unverändert. Die KET-Proteine sind weitaus konservierter als die entsprechenden p53-Proteine vom Menschen und von Ratten, die zu insgesamt 79 % homolog sind. Lediglich im DNA-Bindebereich erreicht ihre Identität 91 %. Menschliches p73 zeigt eine Identität von insgesamt 58 % mit menschlichem KET. Die Konservierung ist wiederum im DNA-Bindebereich mit einer Identität von 86 % am höchsten, während der N-terminale Bereich, mit Ausnahme des Transaktivierungsbereichs, am meisten abweicht. Außer dem Transaktivierungs- und dem DNA-Bindebereich weisen p53, p73 und KET einen gut erhaltenen Oligomerisationsbereich gemeinsam auf. Es ist wahrscheinlich, daß die drei Proteine in der Lage sind, Mischoligomere zu bilden, die spezifische biologische Funktionen haben.

Bei Proteinen, die aus funktionellen Gründen keine Veränderung tolerieren können, wie solche, die Mehrfachbindestellen aufweisen, sind Aminosäuresequenzen im allgemeinen so gut erhalten, wie das bei KET vom Menschen und von Ratten zu beobachten ist. Diese Konservierung läßt vermuten, daß KET ein evolutionäres altes Gen sein kann, das wahrscheinlich bei der Entwicklung und Differenzierung höherer wirbelloser Tiere und Wirbeltiere in die allgemeinen Grundfunktionen einbezogen wurde. p53 kann sich später von seinem Vorläufergen als Protein weiterentwickelt haben, das für spezifische Funktionen wie die Überwachung von Genomschäden verantwortlich ist. Der Umfang der genotoxischen Belastung hängt von physiologischen Faktoren und Umweltfaktoren ab, die, zumindest teilweise bei den verschiedenen Arten unterschiedlich sind. So kann die relative Vielgestaltigkeit von p53, im Vergleich zu KET, die artspezifischen Anforderungen an ein solches System widerspiegeln.

Für das Kartierungsverfahren auf der Grundlage von PCR wurden STS vom Menschen (hKET8) und zwei KET STS von Mäusen (muKET8 und muKET9) amplifiziert (s. Tab. 1).

Die Primerpositionen wurden so definiert, daß Fragmente entstanden sind, die ein Intron, flankiert von Exonsequenzen, enthalten. Das ermöglichte die Identifizierung richtiger PCR-Fragmente durch einen Vergleich mit der bekannten KET-cDNA-Sequenz von Ratten. Speziell für die Kartierung des Ket-Gens von Mäusen haben wurden intronhaltige PCR-Fragmente gewählt, um nach einer Restriktion mit geeigneten Endonucleasen leicht nachweisbare Fragment-Längen-Polymorphismen zu erhalten. Die Exon-Intron-Grenzen wurden durch einen Vergleich der KET-Aminosäuresequenz mit der von p53 und p73 (Schmale und Bamberger, 1997; Kaghad et al, 1997) abgeleitet. Die Exonsequenzen entsprachen der KET-Aminosäuresequenz des Menschen (Abb. 1) wie folgt: hKET9, Aminosäurereste 360-390; muKET8, Aminosäurereste 360 – 400; muKET9, Aminosäurereste 383 – 438. PCRs wurden, wie bereits beschrieben (Lengeling et al, 1995), durchgeführt. Nukleotidsequenzen von Primern für den Mit-Mikrosatellitenmarker D16Mit57 wurden aus der MIT-Mausgenomdatenbank gewonnen. Maussegregationsdaten wurden mit dem GENE-LINK Computerprogramm (Montagutelli, 1990) verarbeitet. hKET8, das Intron 8 umfaßt, wurde mittels der GeneBridge 4 Strahlungshybridkartierungspanels kartiert (Research Genetics, Huntsville, AL). Dieses Panel stellt 91 Strahlungshybridklone des gesamten Humangenoms dar. Das KET-Gen wurde am Humanchromosom 3q27 zwischen den Mikrosatellitenmarkern D3S1580 und D3S1314 kartiert (Abb. 2A). Die wahrscheinlichste Genreihenfolge und -abstände waren D3S1580 – 2,2 cR – WI-6145 – 7,1 cR – KET – 4,7 cR – WI1189 – 8,9 cR – D3S1314. Dieser Bereich ist von *Somatostatin*-, SST (O'Hara et al, 1988) und *Apolipoprotein D*, APOD (Warden et al, 1992) flankiert. Informationen über die chromosomale Lokalisierung von SST und APOD und die Genreihenfolge wurden der Chromosomen 3-Karte entnommen (San Antonio Genome Center; <http://genome.uthcsa.edu/Maps/frame.html>).

3q27 ist der mittlere Teil eines Bereich einer gut dokumentierten Syntenie zum Mauschromosom 16, der sich von *CLCN2* nach *GAP43*, bzw. *Clc2* nach *Gap43* im Mausgenom erstreckt (DeBry und Seldin, 1996; Lengeling et al, 1995). Um

festzustellen, ob der Maus-Ket-Locus in den homologen Bereich fällt, erfolgte eine Kartierung des Ket-Gens unter Verwendung einer Interspeziesrückkreuzung der Maus (C57BL/6J *wrl* + *xSEG/1* +/+)*F₁ *wrl/+* x (C57BL/6J *wrl*+), die ursprünglich für die Kartierung des *Wobbler*-Gens etabliert wurde (Kaupmann et al, 1992). Dieses Interspeziesrückkreuzungspanel wurde für über 150 Loci charakterisiert, die über alle Autosomen und die X-Chromosomen verteilt waren. Es wurde eine verbesserte Karte des Chromosoms 16 für die Kartierung des Chloridkanalgens *Clc2* geschaffen (Lengeling et al, 1995). Beide Maus-KET-PCR-Fragmente lieferten informative Restriktionsfragmentlängenvarianten (RFLVs), die für die Segregationsanalyse verwendet wurden. Das Fragment mit Intron 8 (*muKET8*) wurde mit *Msp1* geschnitten, *muKET9* mit *Rsa1*. KET wurde zwischen *Smst* und *D16Mit63* entdeckt mit Lodscores > 8 (Abb. 2B,C). Das Maushomologe des menschlichen *Apolipoproteins D*, *Apod*, der Genmarker an menschlichem Chr 3q eng verbunden distal zu KET, wurde im *M. musculus* x *M. spretus* Rückkreuzungspanel nicht kartiert, jedoch sind detaillierte Kartierungsdaten verfügbar (Reeves und Cabin, 1997; Warden et al, 1992; Reeves et al, 1997). Es wurde *D16Mit 57* kartiert, das distal zu *Apod* (Reeves und Cabin, 1997) angeordnet ist und einen Polymorphismus von Fragmentlänge zwischen dem *M. musculus* C57BL/6J (111 bp) und *M. spretus* SEG/1 (135 bp) nutzt. Bei 50 Meiosen wurde keine Rekombination von *Ket* und *D16Mit57* festgestellt. Auf der Rückkreuzungstafel von *M. musculus* x *M. spretus* waren die wahrscheinlichsten Genreihenfolge und -abstände *Cen* – *D16Mit87* – 3,9 ± 1,7 cM – *Smst*, *D16Mit102* – 4 ± 2,77 cM – *Ket*, *D16Mit 57* – 14 ± 4,91 cM – *D16Mit63*.

Der Verlust des langen Arms von Chromosom 3 wird selten festgestellt (vgl. Chitayat et al, 1996). Die Symptome mit Eliminierungen 3q27→qter unterscheiden sich erheblich und reichen nicht aus, um ein spezifisches Syndrom abzuleiten. Während in zwei Fällen lediglich kleinere faziale Abnormitäten, Verzögerungen in der Entwicklung und Hypotonien berichtet wurden, zeigten andere ernsthafte, mehrfache kongenitale Abnormitäten, einschließlich Anophtalmie und Hirnathrophie.

In einigen Humankrebsgeweben (z.B. ösophagealem Krebs und squamöse Karzinomen) wurden LOH-Bereiche an Chr 3 entdeckt, die 3q27 enthielten (Sato et al, 1994; Wang et al, 1996). Obwohl statistisch von Bedeutung, war ein Verlust von 3q, der im Vergleich zu anderen chromosomalen Abnormitäten mit einer relativ geringen Häufigkeit auftrat, bei diesen Tumoren zu beobachten.

Vergleichende Kartierungsdaten zeigten Bereiche einer völlig erhaltenen Syntenie zwischen Chr 3q und den Mausechromosomen 3, 9 und 16 (vgl. DeBry und Seldin, 1996), wovon zwei die vorhergesagten Suppressorgene *Loh1* und *Loh2* beherbergen, die auf ausgeprägten Stufen der Tumorentwicklung in einem transgenen Mausmodell des Inselzellkarzinoms (Dietrich et al, 1994; Parangi et al, 1995; Shi et al, 1997) deletiert werden. Das *Ket*-Gen fällt in den gleichen LOH-Bereich mit *Loh2* (LOH etwa 15 cM, 14-29 cM von Cen, flankiert durch die Mikrosatellitenmarker *D16Mit35* und *D16Mit39*; (vgl. Parangi et al, 1995). Von *Loh2* wird angenommen, daß es einen Suppressor der Angiogenese (Parangi et al, 1995) kodiert. Tatsächlich wurde gezeigt, daß das Protein p53 die Angiogenese in Fibroblasten indirekt hemmt durch die positive Regulierung der Thrombospondin-1-Expression (Dameron et al, 1994). Daher ist *Ket* ein Kandidat für *Loh2* durch seine chromosomale Lokalisierung und durch seine putative Funktion.

Tabelle 1

PCR Primer für STS aus den *KET*/*Ket* Regionen

STS	Sequenz (5'→3')	Größe (bp)	Temperierung	Referenz
hKET8	CAGAAAGCAGCAAGTTTCGGAC TGGATGTCATCTGGATACCATG	750	55°C	
muKET8	CAGAAAGCAGCAAGTTTCGGAC AGCTCATCATCTGGGGATCTCC	2.300	65°C	
muKET9	ACACGGAATCCAGATGACTTCC TGCTGCCTGTACGTTTCGATCG	3.100	65°C	
<i>D16Mit57</i>	AAAAAATTTTAAACCATGTGAATGT TGAAGTTTATTATGAGTTGAATCATGC	111 135 ^a	63°C	MIT

Größe des Amplifizierungsproduktes mit C57BL/6J DNA; ^a(SEG/1)

Zitierte Referenzen

- Chitayat, D., Babul, R., Silver, M. M., Jay, V., Teshima, I. E., Babyn, P., and Becker, L. E. (1996). Terminal deletion of the long arm of chromosome 3 [46,XX,del(3)(q27-->qter)]. *Am J Med Genet* **61**, 45-8.
- Dameron, K. M., Volpert, O. V., Tainsky, M. A., and Bouck, N. (1994). Control of angiogenesis in fibroblasts by p53 regulation of thrombospondin-1. *Science* **265**, 1582-4.
- DeBry, R. W., and Seldin, M. F. (1996). Human/mouse homology relationships. *Genomics* **33**, 337-51.
- Dietrich, W., Miller, J., Steen, R., Merchant, M., Damron-Boles, D., Husain, Z., Dredge, R., Daly, M., Ingalls, K., O'Connor, T., and et, a. (1996). A comprehensive genetic map of the mouse genome. *Nature* **380**, 149-52.
- Dietrich, W. F., Radany, E. H., Smith, J. S., Bishop, J. M., Hanahan, D., and Lander, E. S. (1994). Genome-wide search for loss of heterozygosity in transgenic mouse tumors reveals candidate tumor suppressor genes on chromosomes 9 and 16. *Proc Natl Acad Sci U S A* **91**, 9451-5.
- Donehower, L. A., Harvey, M., Slagle, B. L., McArthur, M. J., Montgomery, C. A., Jr., Butel, J. S., and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356**, 215-21.
- Donehower, L. A., Harvey, M., Vogel, H., McArthur, M. J., Montgomery, C. A., Jr., Park, S. H., Thompson, T., Ford, R. J., and Bradley, A. (1995). Effects of genetic background on tumorigenesis in p53-deficient mice. *Mol Carcinog* **14**, 16-22.
- Evans, S. C., and Lozano, G. (1997). The Li-Fraumeni syndrome: an inherited susceptibility to cancer. *Mol Med Today* **3**, 390-5.
- Harvey, M., McArthur, M. J., Montgomery, C. A., Jr., Butel, J. S., Bradley, A., and Donehower, L. A. (1993). Spontaneous and carcinogen-induced tumorigenesis in p53-deficient mice. *Nat Genet* **5**, 225-9.
- Hollstein, M., Sidransky, D., Vogelstein, B., and Harris, C. C. (1991). p53 mutations in human cancers. *Science* **253**, 49-53.

- Kaghad, M., Bonnet, H., Yang, A., Creancier, L., Biscan, J. C., Valent, A., Minty, A., Chalon, P., Lelias, J. M., Dumont, X., Ferrara, P., McKeon, F., and Caput, D. (1997). Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers. *Cell* 90, 809-19.
- Kaupmann, K., Simon-Chazottes, D., Guenet, J., and Jockusch, H. (1992). Wobbler, a mutation affecting motoneuron survival and gonadal functions in the mouse, maps to proximal chromosome 11. *Genomics* 13, 39-43.
- Lengeling, A., Gronemeier, M., Ronsiek, M., Thiemann, A., Jentsch, T. J., and Jockusch, H. (1995). Chloride channel 2 gene (Clc2) maps to chromosome 16 of the mouse, extending a region of conserved synteny with human chromosome 3q. *Genet Res* 66, 175-8.
- Levine, A. J. (1997). p53, the cellular gatekeeper for growth and division. *Cell* 88, 323-31.
- Montagutelli, X. (1990). GENE-LINK: a program in PASCAL for backcross genetic analysis. *J Hered* 81, 490-1.
- O'Hara, B., Bendotti, C., Reeves, R., Oster-Granite, M., Coyle, J., and Gearhart, J. (1988). Genetic mapping and analysis of somatostatin expression in Snell dwarf mice. *Brain Res* 464, 283-92.
- Parangi, S., Dietrich, W., Christofori, G., Lander, E. S., and Hanahan, D. (1995). Tumor suppressor loci on mouse chromosomes 9 and 16 are lost at distinct stages of tumorigenesis in a transgenic model of islet cell carcinoma. *Cancer Res* 55, 6071-6.
- Reeves, R. H., and Cabin, D. E. (1997). Mouse Chromosome 16. *Mamm Genome* 7, 264-73.
- Reeves, R. H., Patch, D., Sharpe, A. H., Borriello, F., Freeman, G. J., Edelhoff, S., and Distech, C. (1997). The costimulatory genes Cd80 and Cd86 are linked on mouse chromosome 16 and human chromosome 3. *Mamm Genome* 8, 581-2.
- Sah, V. P., Attardi, L. D., Mulligan, G. J., Williams, B. O., Bronson, R. T., and Jacks, T. (1995). A subset of p53-deficient embryos exhibit exencephaly. *Nat Genet* 10, 175-80.
- Sato, S., Nakamura, Y., and Tsuchiya, E. (1994). Difference of allelotype between squamous cell carcinoma and adenocarcinoma of the lung. *Cancer Res*

54, 5652-5.

Schmale, H., and Bamberger, C. (1997). A novel protein with strong homology to the tumor suppressor p53. *Oncogene* 15, 1363-7.

Shi, Y. P., Naik, P., Dietrich, W. F., Gray, J. W., Hanahan, D., and Pinkel, D. (1997). DNA copy number changes associated with characteristic LOH in islet cell carcinomas of transgenic mice. *Genes Chromosomes Cancer* 19, 104-11.

Walter, M. A., Spillet, D. J., Thomas, P., Weissenbach, J., and Goodfellow, P. N. (1994). A method for constructing radiation hybrid maps of whole genomes. *Nat Genet* 7, 22-8.

Wang, L., Li, W., Wang, X., Zhang, C., Zhang, T., Mao, X., and Wu, M. (1996). Genetic alterations on chromosomes 3 and 9 of esophageal cancer tissues from China. *Oncogene* 12, 699-703.

Warden, C., Diep, A., Taylor, B., and Lysis, A. (1992). Localization of the gene for apolipoprotein D on mouse chromosome 16. *Genomics* 12, 851-2.

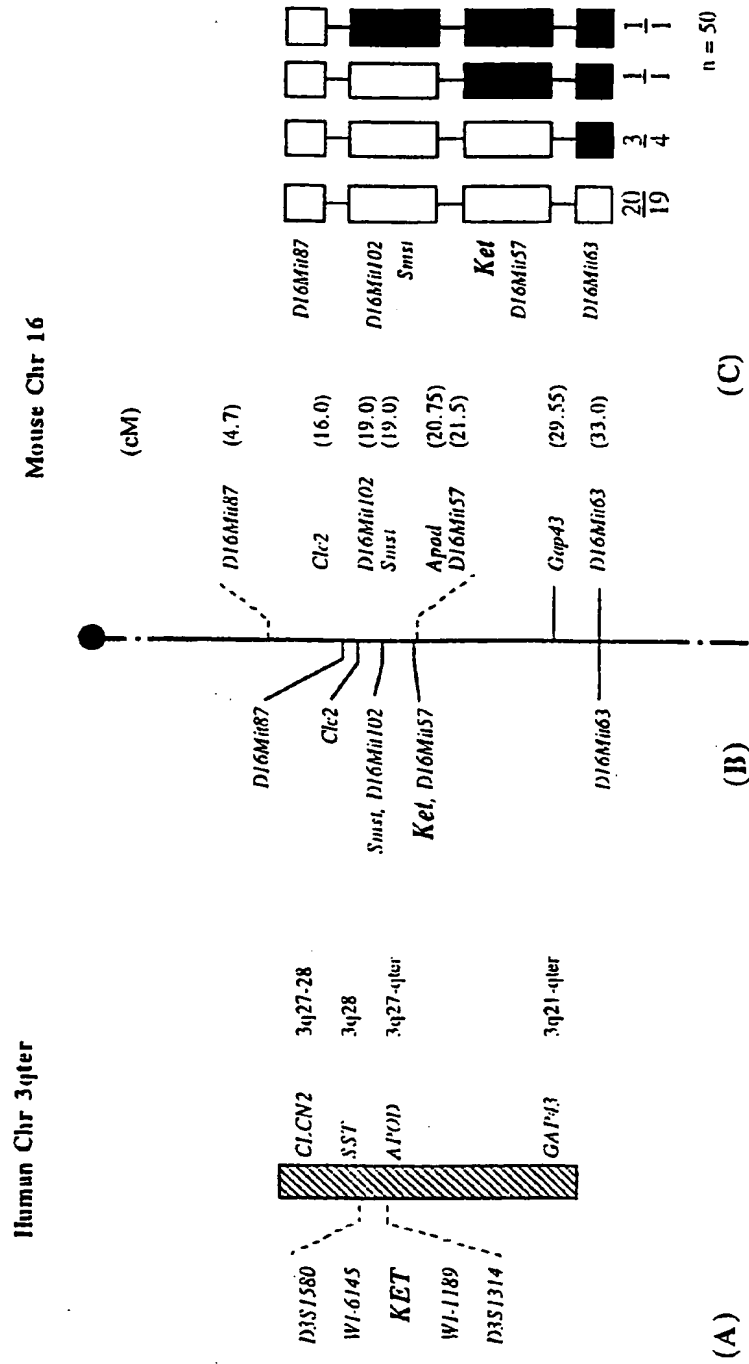
Patentansprüche

1. KET-kodierende Nukleinsäuren, Fragmente, Varianten und Mutationen.
2. KET-Nukleinsäuren nach Anspruch 1, gekennzeichnet durch die KET-cDNA der Ratte mit der Sequenz SEQ ID No. 1 sowie deren Fragmente, Varianten und Mutationen.
3. KET-Nukleinsäuren nach Anspruch 1, gekennzeichnet durch die humane KET-cDNA mit der Sequenz SEQ ID No. 2 sowie deren Fragmente, Varianten und Mutationen.
4. KET-Nukleinsäuren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß T durch U ausgetauscht ist.
5. KET-Nukleinsäuren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß sie vollständig komplementär ist.
6. Polypeptide, für die KET-Nukleinsäuren gemäß einem der Ansprüche 1 bis 5 kodieren.
7. Polypeptide nach Anspruch 6 gekennzeichnet durch die SEQ ID No. 3 sowie deren an einer oder mehreren Stellen durch Austausch von Aminosäuren geänderte Strukturen.
8. Vektoren, die eine KET-Nukleinsäure oder für KET-Polypeptide kodierende DNA nach einem der Ansprüche 1 bis 5 enthalten.
9. Wirtszellen, die die Vektoren gemäß Anspruch 8 enthalten.
10. Verwendung von KET-Nukleinsäure oder Polypeptiden nach einem der Ansprüche 1 bis 7 zum Nachweis von KET-Nukleinsäuren in biologischen Proben.
11. Verwendung nach Anspruch 10, dadurch gekennzeichnet, daß man eine biologische Probe mit mindestens einer Verbindung dieser Nukleinsäuren,

vorzugsweise mit den Verbindungen der SEQ ID No. 1, SEQ ID No. 2 und/oder SEQ ID No. 3, ggf. mit einem Trägermolekül nach an sich üblichen Methoden in Kontakt bringt und der Nachweis anhand des gebildeten Hybridisationskomplexes durch physikalische oder chemische Methoden erfolgt.

12. Verwendung nach Anspruch 10 oder 11, dadurch gekennzeichnet, daß die Nukleinsäure DNA ist, die ggf. eine homozygotische Deletion enthält.
13. Verwendung nach Anspruch 10, dadurch gekennzeichnet, daß die Nukleinsäure RNA ist.
14. Verwendung nach einem der Ansprüche 10 bis 13, dadurch gekennzeichnet, daß die Nukleinsäure markiert ist, vorzugsweise durch ein Radioisotop, eine biolumineszente, eine chemilumineszente oder fluoreszente Verbindung, ein Metallchelat oder ein Enzym.
15. Verwendung nach einem der Ansprüche 10 bis 14, dadurch gekennzeichnet, daß die biologische Probe Tumorgewebe mit erfolgter Angiogenese des Menschen oder der Maus ist.
16. Verwendung von KET-Nukleinsäuren nach einem der Ansprüche 1 bis 5 zum Nachweis der Gegenwart oder Abwesenheit des menschlichen Chromosoms 3q27 oder dessen Fragmenten anhand des Hybridationsproduktes zwischen chromsomal DNA und der KET-Nukleinsäure.
17. Verwendung von KET-Nukleinsäuren nach einem der Ansprüche 1 bis 5 zum Nachweis der Gegenwart oder Abwesenheit des Mauschromosoms 16 oder dessen Fragmenten anhand des Hybridationsproduktes zwischen chromsomal DNA und der KET-Nukleinsäure.
18. Testkit zum Nachweis oder Veränderungen von KET-Nukleinsäuren enthaltend
 - mindestens eine KET-Nukleinsäure oder ein Polypeptid gemäß einem der Ansprüche 1 bis 7 oder eine Hybridisationsprobe.

1



A66 2

SEQUENZPROTOKOLL

(1) ALLGEMEINE INFORMATION:

(i) ANMELDER:

- (A) NAME: Fraunhofer-Gesellschaft zur Förderung der
angewandten Forschung e.V.
- (B) STRASSE: Leonrodstr. 68
- (C) ORT: München
- (E) LAND: Deutschland
- (F) POSTLEITZAHL: D-80636

(ii) ANMELDETITEL: Tumorsuppressorgen der p53-Familie

(iii) ANZAHL DER SEQUENZEN: 3

(iv) COMPUTER-LESBARE FORM:

- (A) DATENTRÄGER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)

(2) INFORMATION ZU SEQ ID NO: 1:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 4708 Basenpaare
- (B) ART: Nukleinsäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNS

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

TAAGTGAGTT CCTCAGCCCA GAGGTGTTCC AGCATATCTG GGATTTTCTG GAACAGCCTA	60
TATGCTCAGT ACAGCCCATC GACTTGAAC TGTGGACGA ACCATCAGAA AATGGTGCAA	120
CAAACAAGAT TGAGATTAGC ATGGATTGTA TCCGCATGCA AGACTCAGAC CTCAGTGACC	180
CCATGTGGCC ACAGTACACG AACCTGGGGC TCCTGAACGG CATGGACCAG CAGATTGAGA	240
ACGGCTCCTC ATCTACCAGC CCCTATAACA CAGACCATGC ACAGAACAGC GTGACGGCAC	300
CTCTCGCCCTA TGCACAGCCC AGCTCAACCT TCGATGCCCT TTCTCCATCC CTTGCCATTC	360
CCTCCAACAC AGATTACCCA GGCCACACA GCTTCGATGT GTCCTTCCAG CAGTCAAGCA	420
CCGCCAAGTC AGCTACCTGG ACGTATTCCA CCGAACTGAA GAACTCTAC TGCCAGATTG	480
CAAAGACCTG CCCCATCCAG ATCAAGGTGA TGACCCACC CCCACAGGGC GCCGTCATTC	540
GTGCCATGCC TGTCTACAAG AAAGCCGAGC ATGTCACCGA GGTGTGAAA CGATGTCCTA	600
ACCACGAGCT GAGCCGCGAG TTCAATGAGG GACAGATTGC CCCTCCCACT CATCTGATTC	660

GAGTAGAAGG GAACAGCCAT GCCCAGTATG TAGAAGATCC TATCACAGGA AGGCAGAGCG	720
TGCTGGTCCC TTATGAGCCA CCACAGGTTG GCACTGAATT CACAACAGTC CTGTACAATT	780
TCATGTGCAA CAGCAGCTGT GTCGGAGGAA TGAACCGCCG TCCAATTTTA ATCATCGTTA	840
CTCTGGAAAC CAGAGATGGG CAAGTCCTGG GCCGACGTTG CTTTGAGGCC CGGATCTGCG	900
CTTGCCCAGG AAGAGACCGG AAGGCCGATG AAGACAGCAT CAGAAAGCAG CAAGTATCAG	960
ACAGCGCAAA GAACGGCGAT GGTACGAAGC GCCCTTTCCG TCAGAATACC CACGGAATCC	1020
AGATGACTTC CATCAAGAAA CGGAGATCCC CAGATGATGA GCTGCTGTAC CTACCAGTGA	1080
GAGGCCGTGA GACTTATGAA ATGCTGCTCA AGATCAAGGA GTCGCTCGAG CTCATGCAGT	1140
ATCTCCCTCA GCACACGATC GAGACGTACA GGCAGCAGCA GCAGCAGCAG CACCAACACC	1200
TACTTCAGAA ACAGACCTCG ATGCAGTCTC AGTCTTCATA CGGTAACAGC TCACCACCTC	1260
TGAACAAAAT GAACAGCATG AACAAGCTGC CGTCTGTGAG CCAGCTTATC AACCACAGC	1320
AGCGCAACGC CCTGACTCCC ACCACCATGC CTGAGGGCAT GGGAGCCAAC ATTCTATGA	1380
TGGGCACTCA CATGCCAATG GCTGGAGACA TGAATGGACT CAGCCCCACC CAAGCTCTTC	1440
CTCCTCCACT CTCCATGCCC TCCACCTCCC ACTGCACCCC CCCACCTCCG TACCCAACAG	1500
ACTGCAGCAT TGTCAGTTT TTAGCAAGGT TGGGCTGTTT ATCATGTCTG GACTATTTC	1560
CGACCCAGGG GCTGACCACC ATCTATCAGA TTGAGCATTG CTCCATGGAT GATTTGGCAA	1620
GTCTGAAGAT CCCTGAGCAG TTCCGACATG CCATCTGGAA GGGGATCCTG GACCACAGGC	1680
AGCTGCATGA CTTCTCCTCA CCTCCGCATC TCCTGAGAAC CCCCAGTGGT GCCTCTACAG	1740
TCAGTGTTGG CTCCAGTGAG ACCCGTGGAG AACGTGTGAT TGATGCCGTG CGCTTTACTC	1800
TCCGCCAGAC CATCTCTTTC CCACCCCGTG ATGAGTGGAA CGATTTC AAC TTTGACATGG	1860
ATTCCCGTCG CAACAAGCAG CAGCGCATCA AAGAGGAAGG AGAATGAACG TCCGTCGCCG	1920
GGTTCTTCCT GTTTTCTTCC TCCTCCCAGC TCCCACAGGG CACGCCTGCT TGATCCTCAA	1980
AGCCTTCTCG CTAGCTCTCC TCCTCCTCCT TCTCAGTCTG GTTTCTAAAG GGACGGAGAA	2040
TTAAGAGGCT ACCTGTTACC TAAAGTCTGA CCTGTACCT GATTCTGATC CTGGCTTTAA	2100
GCCTTCAATA CTCTTGCTTG CAAGATGCGT TGACATTGCT AGATAGACGT TAGCAGAGAA	2160
GCAGTGGGTC TCTCTAAGCA CTGGAGATCG CTCATTGACT TTTATAAAGC ATTTTCAGCC	2220
TTATAGTCTA AGACTATATA TATAAATATA TAAATATACA ATATATATTT CGGGTGGGGG	2280
TATTGAGTAT TGTTTAAATG TAATTTAATG GAAATCGAGT TGCACTTATC AACCTTCTTT	2340
GGAAATTTGCT TGTTTGGTT GGCTGATCTG TACCCCTTTC TCAGGGGTAT CATGTATGGT	2400

GACAGATATT TAGAGTTGAA TGGTCTATGT GAGTAACAGT GATATATAGG TCCTCTCCTT	2460
TCTTTGGATG ATTGCCGTTT AGCACATCAA ACCTGTGGAT GCGTCCAGTC TGTTTACCAT	2520
TGCTCCTTAT GAGGTAAAAC TGCATATACT GTCAGTCTAT TTTATGTTAC TGGTGTCCAT	2580
TCCAGTTAGG CTGGTTCAC TGTGGCCAT TCCAAGCAAA ATTTTATGTT TGCTTTGTCA	2640
CACACTAGAA GACAGGGCAT CATCTCTTGC TTTTGTGTTGA GAATGAGGAG TACTTTTTTT	2700
TTTTTCTGGA AAATCTTAAA TGGTCCAAAT CAGCCATTCC AAATGGCTGA TGAAATGTAG	2760
CCAATATAGC AGTTAGCTCT CTAAAATTTA AGACCCAACA CCCTCGTATT TATTAGTAAA	2820
ACAAAAATGA AACATTTGCT GTCATTAGAG TAGCCTTAAA ATTAAATTC AATACCAGAT	2880
TGACTGAGTA AACTATGCAT TCAATGTTGT TGTGAGAATT GGGGCTAATT AGTCAGGATG	2940
ATTGGAATTT GTGTAGTTTT TTATGGTGAG TTGCAATATC TATTTAGGAA GGTTTCAGGAA	3000
TAATAAGAAT GACTCAGAAA TACTCAATCT CCGTGACAAC AGAAAGCAAT CTCACCAAAC	3060
TCTGAATTTA AACCCCTTTT GAAACATGGA GTGAGGCTTG GGAAATGTAC CTTTTAAAGA	3120
CTTTCCTATC TATAAGACAC TGCATGCAGG GGCAAGTTTA ATCTCTCATC AAGGTGGAAA	3180
ATAAGAATAG TAGCTCGGAA ACTACAAACT TGCTAGTGTA GCTTTCACAT GGCATGAGCT	3240
CAACTATTGT TATTTTCCTC TTTATCATCA AAGCTCCATT GCTGTAGAAA GCAGAGGTGA	3300
AGACCCAGTT TTCCACCTGA CACTTTCGGG GCAAGGCATA GACCAAGAAC TGTCTACAAA	3360
ACCAGGGCAA AGCTCTTCAG TGAAGCTGTT TAATTCACAT GGAGAAACAC TTGTTTCCCA	3420
CTTTGGGAAA GCATGCAACA GTGTTCCCCC TAGATGTTTT GGAAACATTT TGAGTCAAAAT	3480
ATATTTTTCC CAGACTAAAC CAGGCTAATG AGCTCTACAA TCCTCCTGCA CATTTTGGTA	3540
AAGGGCTGTC ATTGCACAGG AGCTCCCAT TTTATCTTAA AGTGCAAATG GGCTAATACG	3600
CCTACGAAAT GTAATGTATG GGTTTTGCCA GAAAATAGTA TATTGTGTAC ACGTGTCTGT	3660
GTGTGAGTGT GAGAGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGAAA TTGCATACTA	3720
TGCTGGTTTT GTTTGTTACT CTTTCTCTTG GGGATAGTTG GGTTTCCAG AACCACAGAC	3780
GAACTTTTTT TTTGTTGCTG TTTTATATT TTTGCAGAAA CACCATTAG TGAGAATTCA	3840
ATGTCAAATT AGACATGACA CCTTAATTGT AAGAAGGGGG GAGAGGGAAA GTTGGTTTTT	3900
TTTAATTTTT TAAATTTTG TATACTAAAG AGAATGAGTC CTTAATTTCA ACATTCTGTT	3960
GCATTTAAAT AATGATAAGC ATCATTAAC TCTGTAACAA CTTCCAGCT TGGCAAATTC	4020
AATGCATGGA GAACAAAGCT GGGCCTTAGC CATGTTAGGG AGAAAAATGG CTTCTTGGGG	4080

```

GTTGTGAGCA TTTGGGTTGC TTTAGCACCG TTGAGGTGGC ACAGGGGACT CCTGAGGCAT 4140
TTCAGCACTA CTTACGTAGC ACTAGGGACT CGGAAATTCC TGTACTGTAG CTAATGATTT 4200
TGGCGTTCAC CATTAGCAGT AGATAGGCCG TTTCTCTCCT CACACCAGTG TTAAGCGTGT 4260
GAGTAGCCAG AGCTGTGGGG AAGAGCATGG AGAACAGACG TCTGCTGGAT GCCTCTCACC 4320
GGAGAATGAG ATTCCTTCGC GTGGTGGTGA AGTAGGATAG GAAGCAGGAG TCTCCTTGTT 4380
AGTCCAGTTA GCTATTGTTT TCTTGATATT CCCCCC AAA ACATTGACTA TGAGAGATAT 4440
GTGGGGCTTT TTTATTTT TAATTGTACA AAATTAAACA AATATGAAAT GTTTTATATA 4500
CTTTATTAAT GTTTTTTTT AAAAGGTA CTCTTATAGA CATGATCCTT TTTTACAGG 4560
TTCAGTTGCT TGCCCTTGG TATTTTGTG TTATGGGCTA TGGTGAGCCT GAGGCAAATC 4620
TATAAGCCAT TTTTGTGTC CAGGACATGC AATAAAATTT AAAAATAAAT GAAATAACAC 4680
TGAAAAAAAA AAAAAAAAAA AAAAAAAAA 4708

```

(2) INFORMATION ZU SEQ ID NO: 2:

- (i) SEQUENZ CHARAKTERISTIKA:
- (A) LÄNGE: 4846 Basenpaare
 - (B) ART: Nukleinsäure
 - (C) STRANGFORM: Einzel
 - (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNS

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

```

CGTTGATATC AAAGACAGTT GAAGGAAATG AATTTTGAAA CTTACGGTG TGCCACCCTA 60
CAGTACTGCC CTGACCCTTA CATCCAGCGT TTCGTAGAAA CCCCAGCTCA TTTCTCTTGG 120
AAAGAAAGTT ATTACCGATC CACCATGTCC CAGAGCACAC AGACAAATGA ATTCCTCAGT 180
CCAGAGGTTT TCCAGCATAT CTGGGATTTT CTGGAACAGC CTATATGTTT AGTTCAGCCC 240
ATTGACTTGA ACTTTGTGGA TGAACCATCA GAAGATGGTG CGACAAACAA GATTGAGATT 300
AGCATGGACT GTATCCGCAT GCAGGACTCG GACCTGAGTG ACCCATGTG GCCACAGTAC 360
ACGAACCTGG GGCTCCTGAA CAGCATGGAC CAGCAGATTC AGAACGGCTC CTCGTCCACC 420
AGTCCCTATA ACACAGACCA CGCGCAGAAC AGCGTCACGG CGCCCTCGCC CTACGCACAG 480
CCCAGCTCCA CCTTCGATGC TCTCTCTCCA TCACCCGCCA TCCCCTCCAA CACCGACTAC 540
CCAGGCCCGC ACAGTTTCGA CGTGTCCTTC CAGCAGTCGA GCACCGCCAA GTCGGCCACC 600
TGGACGTATT CCACTGAACT GAAGAACTC TACTGCCAAA TTGCAAAGAC ATGCCCCATC 660

```

CAGATCAAGG TGATGACCCC ACCTCCTCAG GGAGCTGTTA TCCGCGCCAT GCCTGTCTAC	720
AAAAAAGCTG AGCACGTCAC GGAGGTGGTG AAGCGGTGCC CCAACCATGA GCTGAGCCGT	780
GAATTCAACG AGGGACAGAT TGCCCCCTCCT AGTCATTTGA TTCGAGTAGA GGGGAACAGC	840
CATGCCCAGT ATGTAGAAGA TCCCATCACA GGAAGACAGA GTGTGCTGGT ACCTTATGAG	900
CCACCCCAGG TTGGCACTGA ATTCACGACA GTCTTGTAACA ATTCATGTG TAACAGCAGT	960
TGTGTTGGAG GGATGAACCG CCGTCCAATT TTAATCATTG TTAATCTGGA AACCAGAGAT	1020
GGGCAAGTCC TGGGCCGACG CTGCTTTGAG GCGCGGATCT GTGCTTGCCC AGGAAGAGAC	1080
AGGAAGGCGG ATGAAGATAG CATCAGAAAG CAGCAAGTTT CGGACAGTAC AAAGAACGGT	1140
GATGGTACGA AGCGCCCGTT TCGTCAGAAC ACACATGGTA TCCAGATGAC ATCCATCAAG	1200
AAACGAAGAT CCCCAGATGA TGAAGTGTG TACTTACCAG TGAGGGGCGG TGAGACTTAT	1260
GAAATGCTGT TGAAGATCAA AGAGTCCCTG GAACTCATGC AGTACCTTCC TCAGCACACA	1320
ATTGAAACGT ACAGGCAACA GCAACAGCAG CAGCACCAGC ACTTACTTCA GAAACAGACC	1380
TCAATACAGT CTCCATCTTC ATATGGTAAC AGCTCCCCAC CTCTGAACAA AATGAACAGC	1440
ATGAACAAGC TGCCTTCTGT GAGCCAGCTT ATCAACCCCTC AGCAGCGCAA CGCCCTCACT	1500
CCTACAACCA TTCCTGATGG CATGGGAGCC AACATTCCCA TGATGGGCAC CCACATGCCA	1560
ATGGCTGGAG ACATGAATGG ACTCAGCCCC ACCCAGGCAC TCCCTCCCC ACTCTCCATG	1620
CCATCCACCT CCCAGTGCAC ACCCCCACCT CCGTATCCCA CAGATTGCAG CATTGTCAGT	1680
TTCTTAGCGA GGTGGGCTG TTCATCATGT CTGGACTATT TCACGACCCA GGGGCTGACC	1740
ACCATCTATC AGATTGAGCA TTAATCCATG GATGATCTGG CAAGTCTGAA AATCCCTGAG	1800
CAATTTGAC ATGCGATCTG GAAGGGCATC CTGGACCACC GGCAGCTCCA CGAATTCTCC	1860
TCCCCTTCTC ATCTCCTGCG GACCCCAAGC AGTGCCTCTA CAGTCAGTGT GGGCTCCAGT	1920
GAGACCCGGG GTGAGCGTGT TATTGATGCT GTGCGATTCA CCCTCCGCCA GACCATCTCT	1980
TTCCACCCC GAGATGAGTG GAATGACTTC AACTTTGACA TGGATGCTCG CCGCAATAAG	2040
CAACAGCGCA TCAAAGAGGA GGGGGAGTGA GCCTCACCAT GTGAGCTCTT CCTATCCCTC	2100
TCCTAACTGC CAGCNCCTTA AAAGCACTCC TGCTTAATCT TCAAAGCCTT CTCCTAGCT	2160
CCTCCCCTTC CTCTTGCTG ATTTCTTAGG GGAAGGAGAA GTAAGAGGCT ACCTCTTACC	2220
TAACATCTGA CCTGGCATCT AATTCTGATT CTGGCTTTAA GCCTTCAAAA CTATAGCTTG	2280
CAGAAGTGA GCTGCCATGG CTAGGTAGAA GTGAGCAAAA AAGAGTTGGG TGTCTCCTTA	2340
AGCTGCAGAG ATTTCTCATT GACTTTTATA AAGCATGTTT ACCCTTATAG TCTAAGACTA	2400

TATATATAAA TGTATAAATA TACAGTATAG ATTTTGGGTG GGGGGGCATT GAGTATTGTT	2460
TAAAATGTAA TTAAATGAA AGAAGATTGA GTTGCACTTA TTGACCATTT TTTAATTTAC	2520
TTGTTTTGGA TGGCTTGTCT ATACTCCTTC CCTTAAGGGG TATCATGTAT GGTGATAGGT	2580
ATCTAGAGCT TAATGCTACA TGTGAGTGAC GATGATGTAC AGATTCTTTC AGTTCTTTGG	2640
ATTCTAAATA CATGCCACAT CAAACCTTTG AGTAGATCCA TTTCCATTGC TTATTATGTA	2700
GGTAAGACTG TAGATATGTA TTCTTTTCTC AGTGTGGTA TATTTTATAT TACTGACATT	2760
TCTTCTAGTG ATGATGGTTC ACGTTGGGGT GATTTAATCC AGTTATAAGA AGAAGTTCAT	2820
GTCCAAACGT CCTCTTTAGT TTTTGGTTGG GAATGAGGAA AATTCTTAAA AGGCCCATAG	2880
CAGCCAGTTC AAAAACACCC GACGTCATGT ATTTGCGCAT ATCAGTAACC CCCTTAAATT	2940
TAATACCAGA TACCTTATCT TACAATATTG ATTGGGAAAA CATTGCTGC CATTACAGAG	3000
GTATTAAAAC TAAATTTTAC TACTAGATTG ACTAACTCAA ATACACATTT GCTACTGTTG	3060
TAAGAATTCT GATTGATTTG ATTGGGATGA ATGCCATCTA TCTAGTTCTA ACAGTGAAGT	3120
TTTACTGTCT ATTAATATTC AGGGTAAATA GGAATCATTC AGAAATGTTG AGTCTGTACT	3180
AAACAGTAAG ATATCTCAAT GAACCATAAA TTCAACTTTG TAAAAATCTT TTGAAGCATA	3240
GATAATATTG TTTGGTAAAT GTTCTTTTGG TTTGGTAAAT GTTCTTTTAA AAGACCCTCC	3300
TATTCTATAA AACTCTGCAT GTAGAGGCTT GTTTACCTTT CTCTCTCTAA GGTTTACAAT	3360
AGGAGTGGTG ATTTGAAAAA TATAAAATTA TGAGATTGGT TTTCTGTGG CATAAATTGC	3420
ATCACTGTAT CATTTTCTTT TTTAACCGGT AAGAGTTTCA GTTGTGTTGA AAGTAAGTGT	3480
GAGAACCCAG TTTCCCGTCC ATCTCCCTTA GGGACTACCC ATAGACATGA AAGGTCCCCA	3540
CAGAGCAAGA GATAAGTCTT TCATGGCTGC TGTGCTTAA ACCACTTAAA CGAAGAGTTC	3600
CCTTGAAACT TTGGGAAAAC ATGTTAATGA CAATATTCCA GATCTTTCAG AAATATAACA	3660
CATTTTTTTG CATGCATGCA AATGAGCTCT GAAATCTTCC CATGCATTCT GGTCAAGGGC	3720
TGTCATTGCA CATAAGCTTC CATTTTAATT TTAAAGTGCA AAAGGGCCAG CGTGGCTCTA	3780
AAAGGTAATG TGTGGATTGC CTCTGAAAAG TGTGTATATA TTTTGTGTGA AATTGCATAC	3840
TTTGTATTTT GATTATTTTT TTTTCTTCT TGGGATAGTG GGATTTCAG AACCACACTT	3900
GAAACCTTTT TTTATCGTTT TGTATTTTC ATGAAAATAC CATTTAGTAA GAATACCACA	3960
TCAAATAAGA AATAATGCTA CAATTTTAAG AGGGGAGGGA AGGGAAAGTT TTTTTTTATT	4020
ATTTTTTTAA AATTTTGTAT GTTAAAGAGA ATGAGTCCTT GATTTCAAAG TTTTGTGTGA	4080

```

CTTAAATGGT AATAAGCACT GTAAACTTCT GCAACAAGCA TGCAGCTTTG CAAACCCATT 4140
AAGGGGAAGA ATGAAAGCTG TTCCTTGGTC CTAGTAAGAA GACAAACTGC TTCCCTTACT 4200
TTGCTGAGGG TTTGAATAAA CCTAGGACTT CCGAGCTATG TCAGTACTAT TCAGGTAACA 4260
CTAGGGCCTT GGAAATTCCT GTACTGTGTC TCATGGATTT GGCCTAGCC AAAGCGAGGC 4320
ACCTTACTG GCTTACCTCC TCATGGCAGC CTACTCTCCT TGAGTGTATG AGTAGCCAGG 4380
GTAAGGGGTA AAAGGATAGT AAGCATAGAA ACCACTAGAA AGTGGGCTTA ATGGAGTTCT 4440
TGTGGCCTCA GCTCAATGCA GTTAGCTGAA GAATTGAAAA GTTTTGTGTT GGAGACGTTT 4500
ATAAACAGAA ATGGAAAGCA GAGTTTTCAT TAAATCCTTT TACCTTTTTT TTTTCTTGGT 4560
AATCCCCTAA AATAACAGTA TGTGGGATAT TGAATGTAA AGGGATATTT TTTTCTATT 4620
ATTTTTATAA TTGTACAAAA TTAAGCAAAT GTTAAAAGTT TTATATGCTT TATTAATGTT 4680
TTCAAAGGT ATTATACATG TGATACATTT TTTAAGCTTC AGTTGCTTGT CTTCTGGTAC 4740
TTTCTGTTAT GGGCTTTTGG GGAGCCAGAA GCCAATCTAC AATCTCTTTT TGTTTGCCAG 4800
GACATGCAAT AAAATTTAAA AAATAAATAA AACTAATTA AGAAAT 4846

```

(2) INFORMATION ZU SEQ ID NO: 3:

(i) SEQUENZ CHARAKTERISTIKA:

- (A) LÄNGE: 680 Aminosäuren
- (B) ART: Aminosäure
- (C) STRANGFORM: Einzel
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

```

Met Asn Phe Glu Thr Ser Arg Cys Ala Thr Leu Gln Tyr Cys Pro Asp
1           5           10           15
Pro Tyr Ile Gln Arg Phe Val Glu Thr Pro Ala His Phe Ser Trp Lys
20          25          30
Glu Ser Tyr Tyr Arg Ser Thr Met Ser Gln Ser Thr Gln Thr Asn Glu
35          40          45
Phe Leu Ser Pro Glu Val Phe Gln His Ile Trp Asp Phe Leu Glu Gln
50          55          60
Pro Ile Cys Ser Val Gln Pro Ile Asp Leu Asn Phe Val Asp Glu Pro
65          70          75          80
Ser Glu Asp Gly Ala Thr Asn Lys Ile Glu Ile Ser Met Asp Cys Ile
85          90          95

```

Arg Met Gln Asp Ser Asp Leu Ser Asp Pro Met Trp Pro Gln Tyr Thr
 100 105 110
 Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Gln Asn Gly Ser
 115 120 125
 Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser Val Thr
 130 135 140
 Ala Pro Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Ala Leu Ser
 145 150 155 160
 Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro His Ser
 165 170 175
 Phe Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala Thr Trp
 180 185 190
 Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala Lys Thr
 195 200 205
 Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Pro Gln Gly Ala Val
 210 215 220
 Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr Glu Val
 225 230 235 240
 Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn Glu Gly
 245 250 255
 Gln Ile Ala Pro Pro Ser His Leu Ile Arg Val Glu Gly Asn Ser His
 260 265 270
 Ala Gln Tyr Val Glu Asp Pro Ile Thr Gly Arg Gln Ser Val Leu Val
 275 280 285
 Pro Tyr Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Val Leu Tyr
 290 295 300
 Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg Arg Pro
 305 310 315 320
 Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val Leu Gly
 325 330 335
 Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg Asp Arg
 340 345 350
 Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp Ser Thr
 355 360 365
 Lys Asn Gly Asp Gly Thr Lys Arg Pro Phe Arg Gln Asn Thr His Gly
 370 375 380
 Ile Gln Met Thr Ser Ile Lys Lys Arg Arg Ser Pro Asp Asp Glu Leu
 385 390 395 400

Leu Tyr Leu Pro Val Arg Gly Arg Glu Thr Tyr Glu Met Leu Leu Lys
 405 410 415
 Ile Lys Glu Ser Leu Glu Leu Met Gln Tyr Leu Pro Gln His Thr Ile
 420 425 430
 Glu Thr Tyr Arg Gln Gln Gln Gln Gln His Gln His Leu Leu Gln
 435 440 445
 Lys Gln Thr Ser Ile Gln Ser Pro Ser Ser Tyr Gly Asn Ser Ser Pro
 450 455 460
 Pro Leu Asn Lys Met Asn Ser Met Asn Lys Leu Pro Ser Val Ser Gln
 465 470 475 480
 Leu Ile Asn Pro Gln Gln Arg Asn Ala Leu Thr Pro Thr Thr Ile Pro
 485 490 495
 Asp Gly Met Gly Ala Asn Ile Pro Met Met Gly Thr His Met Pro Met
 500 505 510
 Ala Gly Asp Met Asn Gly Leu Ser Pro Thr Gln Ala Leu Pro Pro Pro
 515 520 525
 Leu Ser Met Pro Ser Thr Ser Gln Cys Thr Pro Pro Pro Pro Tyr Pro
 530 535 540
 Thr Asp Cys Ser Ile Val Ser Phe Leu Ala Arg Leu Gly Cys Ser Ser
 545 550 555 560
 Cys Leu Asp Tyr Phe Thr Thr Gln Gly Leu Thr Thr Ile Tyr Gln Ile
 565 570 575
 Glu His Tyr Ser Met Asp Asp Leu Ala Ser Leu Lys Ile Pro Glu Gln
 580 585 590
 Phe Arg His Ala Ile Trp Lys Gly Ile Leu Asp His Arg Gln Leu His
 595 600 605
 Glu Phe Ser Ser Pro Ser His Leu Leu Arg Thr Pro Ser Ser Ala Ser
 610 615 620
 Thr Val Ser Val Gly Ser Ser Glu Thr Arg Gly Glu Arg Val Ile Asp
 625 630 635 640
 Ala Val Arg Phe Thr Leu Arg Gln Thr Ile Ser Phe Pro Pro Arg Asp
 645 650 655
 Glu Trp Asn Asp Phe Asn Phe Asp Met Asp Ala Arg Arg Asn Lys Gln
 660 665 670
 Gln Arg Ile Lys Glu Glu Gly Glu
 675 680